Fluid Flow Loading Analysis of Packet Switching Networks

Jonathan S. Turner

Recent research in switching has concentrated on various forms of statistical switching networks capable of supporting user connections of arbitrary bandwidth. Fast packet switching is one approach that has gained a lot of attention and is being studied by many researchers. This paper addresses the problem of how the configuration of user connections affects the loading of the internal links of a switching fabric used for fast packet switching. It introduces a systematic method of analyzing the effects of a given traffic configuration and applies this method to the analysis of several proposed switching architectures. While the method does...
Read complete abstract on page 2.

Recommended Citation

Turner, Jonathan S., "Fluid Flow Loading Analysis of Packet Switching Networks" Report Number: WUCS-87-16 (1987). *All Computer Science and Engineering Research.*
https://openscholarship.wustl.edu/cse_research/801

Fluid Flow Loading Analysis of Packet Switching Networks

Jonathan S. Turner

Complete Abstract:

Recent research in switching has concentrated on various forms of statistical switching networks capable of supporting user connections of arbitrary bandwidth. Fast packet switching is one approach that has gained a lot of attention and is being studied by many researchers. This paper addresses the problem of how the configuration of user connections affects the loading of the internal links of a switching fabric used for fast packet switching. It introduces a systematic method of analyzing the effects of a given traffic configuration and applies this method to the analysis of several proposed switching architectures. While the method does not completely characterized the performance of these networks, it does provide a useful first-order characterization and is a valuable adjunct to queueing and simulation models.

FLUID FLOW LOADING ANALYSIS OF
PACKET SWITCHING NETWORKS

Jonathan S. Turner

WUCS-87-16

Department of Computer Science
Washington University
Campus Box 1045
One Brookings Drive
Saint Louis, MO 63130-4899

Abstract

Recent research in switching has concentrated on various forms of statistical switching networks capable of supporting user connections of arbitrary bandwidth. Fast packet switching is one approach that has gained a lot of attention and is being studied by many researchers. This paper addresses the problem of how the configuration of user connections affects the loading of the internal links of a switching fabric used for fast packet switching. It introduces a systematic method of analyzing the effects of a given traffic configuration and applies this method to the analysis of several proposed switching architectures. While the method does not completely characterize the performance of these networks, it does provide a useful first-order characterization and is a valuable adjunct to queueing and simulation models.

Fluid Flow Loading Analysis of Packet Switching Networks

Jonathan S. Turner

1. Introduction

Fast packet switching is the focus of much of the current research in advanced switching architectures. Kulzer and Montgomery [17] describe a system under development at AT&T Bell Laboratories since about 1981. During that same period a group of researchers at CNET in France has developed a system called Prelude, based on fast packet switching [5]. More recently groups at Bell Communications Research [13], Bell Telephone Manufacturing [6], CSELT in Italy [11], NEC in Japan [19] and Washington University [21] have undertaken similar research efforts.

Researchers have studied various types of switching fabrics to support a fast packet network. One class of fabrics is based on the delta networks, a well-known class of fabrics, first proposed as interconnection networks in parallel computer systems. Most studies of the performance of such networks focus on the queueing behavior under *uniform random traffic* [7,8,9]. The uniform random traffic assumption is that packets are generated independently at each input and assigned random addresses. While this kind of performance information is important and useful, it does not capture the effects of different loading conditions on the network's behavior. In fast packet switching applications, the load imposed on a switch fabric is the result of user connections that pass through the switch, with each connection entering and leaving on a particular pair of links. Connections remain fixed for fairly long periods of time and may have average data rates that vary over a wide range. In this paper, we focus on the effects of user connections on the loading of links within a switch fabric and develop a systematic method of analyzing the effects of different traffic patterns. Our method allows us to prove theorems characterizing the worst-case loading for various switching fabrics. The paper includes a number of such theorems, both as illustrations of our method and for their inherent interest.

1

Most of the early work on fast packet switching was motivated by the desire to carry both voice and data on a single network supporting point-to-point communication. Recently, the focus has shifted to higher speed applications including video, and interest has been growing in networks that support multipoint communication. Multipoint connections can include an arbitrary number of endpoints and can be used to support video distribution, teleconferencing and LAN interconnection [21]. Switching networks supporting multipoint communication must replicate packets. This leads to new switching structures with new characteristics. The performance of these networks can be studied using queueing models and simulation [2,3]. In this paper we apply our method of loading analysis to characterize the worst-case loading conditions in such networks.

In section 2 of the paper, we introduce our method of analysis and apply it to several switching fabrics for point-to-point communication. These include a standard delta network, a delta network preceded by distribution stages and a variation using four port switch elements, proposed by Lea [16]. In section 3, we apply our method to the analysis of the copy network proposed in [21] and several variations.

2. Networks for Point-to-Point Communication

We define a *packet switching network* (or simply network) as a directed graph $G = (N, L)$ consisting of a set of nodes N and a set of directed arcs or *links* L. In addition, G contains a set of distinguished *input nodes* I and a set of distinguished *output nodes* O. Input and output nodes are also referred to as *ports*. Each input port has a single outgoing link and no incoming links, while each output port contains a single incoming link and no outgoing links.

In this paper, we limit ourselves to networks in which the number of input nodes equals the number of output nodes. When we refer to an n port network, we mean a network with n input nodes and n output nodes, numbered from 0 to $n - 1$. We also limit ourselves to networks, which can be divided into a sequence of stages. We say that input ports are in stage 0 and for $i > 0$, a node v is in stage i if for all links (u, v), u is in stage $i - 1$. A link (u, v) is said to be in stage i if u is in stage i. In the networks we consider, all output ports are in a separate stage by themselves. When we refer to a k stage network, we mean that there are k stages containing internal nodes; that is, we neglect the input and output stages.

When describing particular networks, we will find it convenient to use a composition operation. We denote a composition of two networks X_1 and X_2 by $X_1 \boxdot X_2$, where h is a positive integer. The composition operation yields a new network consisting of one or more copies of X_1 connected to one or more copies
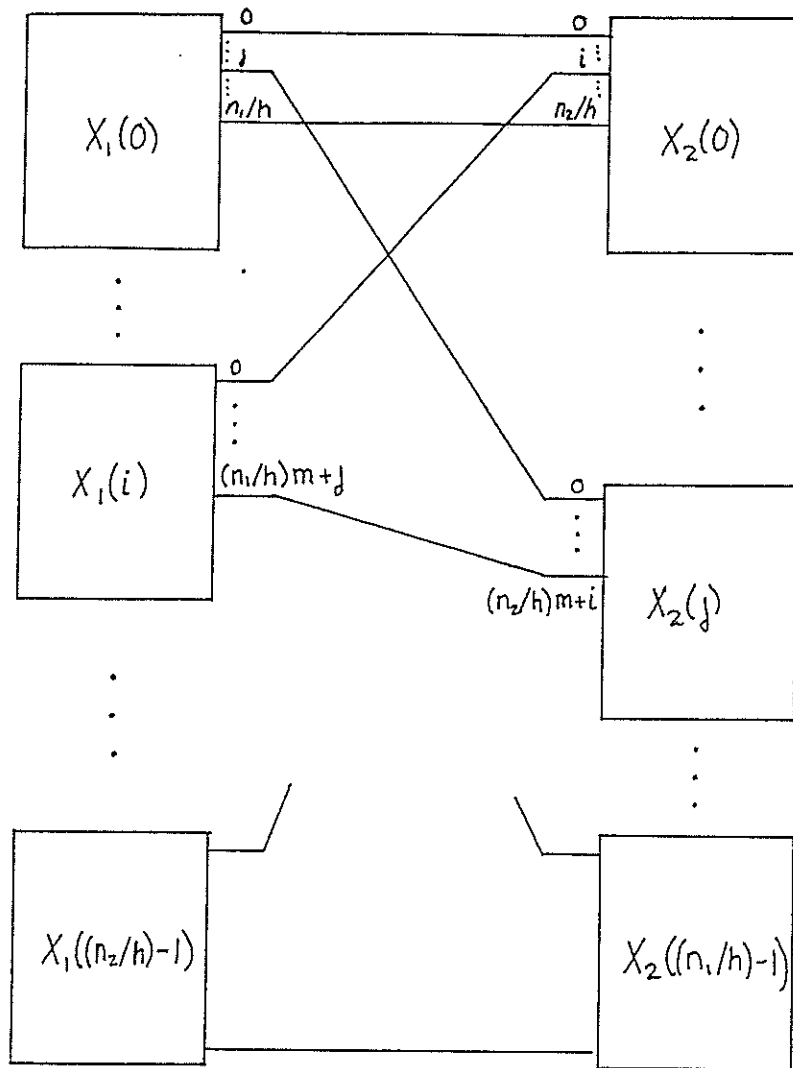
Figure 1: Composition Operation

of X_2, with h links joining each pair of subnetworks. More precisely, if X_1 is an n_1 port network and X_2 is an n_2 port network then $X_1 \boxdot X_2$ is formed by taking n_2/h copies of X_1 numbered from 0 to $(n_2/h) - 1$ followed by n_1/h copies of X_2, numbered from 0 to $(n_1/h) - 1$. Then, for $1 \leq i \leq n_1$, $1 \leq j \leq n_2$, we join $X_1(i)$ to $X_2(j)$ using h links; these links connect output port $(n_1/h)m + j$ of $X_1(i)$ to input port $(n_2/h)m + i$ of $X_2(j)$, where $0 \leq m < h$. Finally, we eliminate the former input and output nodes that are now internal and renumber the input and output nodes of the network as follows; if u was input port i of $X_1(j)$, it becomes input $jn_1 + i$ in the new network; similarly if v was output port i of $X_2(j)$, it becomes output $jn_2 + i$. We also allow composition of more than two networks; the composition $X_1 \boxdot X_2 \boxdot X_3$ is obtained by letting $Y_1 = X_1 \boxdot X_2$ and
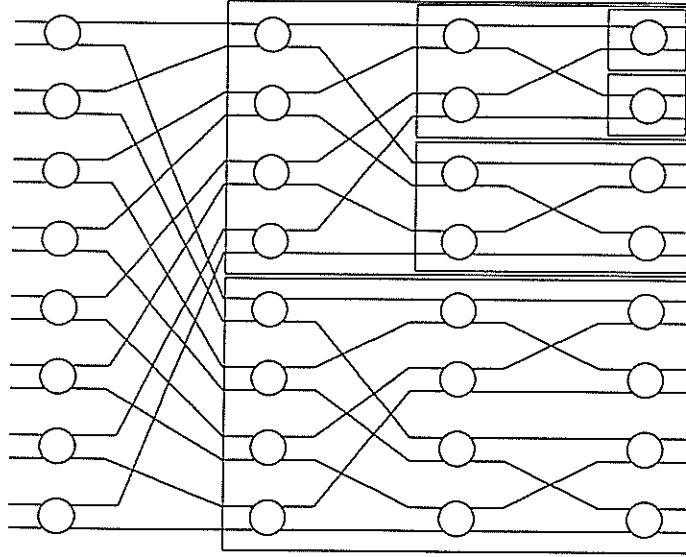
Figure 2: Recursive Construction of Delta Network

$Y_2 = X_2 \boxdot X_3$, then identifying the copies of X_2 in Y_1 and Y_2. This requires of course that the number of copies of X_2 generated by the two initial compositions be the same. Note this is not the same as $(X_1 \boxdot X_2) \boxdot X_3$. The composition operation is illustrated in Figure 1.

A *connection* through a network is defined as a triple (x, y, ρ) where $x \in I$, $y \in O$ and $0 < \rho < 1$. A connection *induces a load* on the various links that lie on paths joining the connection's input and output ports. The load induced by a connection (x, y, ρ) on the link leaving x is defined to be ρ. The magnitude of the induced loads on the internal links depends on the types of the nodes and the topology of the network. In this section, we will consider only a single node type. If α is the sum of the loads induced by a connection (x, y, ρ) on the input links of a node u, and u has i output links that lie on paths from x to y, then the load induced by the connection on each of these output links is α/i and the load induced on all other output links is 0.

A *configuration* is defined as a set of connections. The load induced by a configuration on a link ℓ is simply the sum of the loads induced by the individual connections and is denoted $\lambda_\ell(C)$. A configuration $C = \{c_1 \ldots c_r\}$ is α-bounded if for all input and output links ℓ, $\lambda_\ell(C) \leq \alpha$. We say that a configuration is *legal* if it is 1-bounded and that a network is *robust* if for every legal configuration C, $\lambda_\ell(C) \leq 1$ for *all* links ℓ.

Delta networks form a well-known class of useful switching networks [7,8,9,10].

We can define these recursively using the composition operation. Let D_1 be a network with two input ports and two output ports connected to a single internal node. We then define $D_i = D_1 \Box D_{i-1}$ for all $i \geq 1$. We refer to D_k as a k stage delta network; note that D_k has $n = 2^k$ ports. An example of a 4 stage delta network is given in Figure 2.

Delta networks have been widely studied and have many interesting properties. Most useful is the *self-routing property* that allows paths from inputs to outputs to be easily determined. A related property is that there is a single path connecting any input node to any output node. For the purposes of our loading analysis, this means that a connection (x, y, ρ) induces a load of ρ on all links that lie on that path and a load of 0 on all other links. To illustrate our method of loading analysis, we start with a simple theorem which characterizes the worst-case loading for a delta network.

THEOREM 2.1. *Let $C = \{c_1, \ldots, c_r\}$ be an α-bounded configuration for D_k. Then $\lambda_\ell(C) \leq \alpha\sqrt{n}$ for all links ℓ.*

Proof. Consider a particular link ℓ in stage i. Notice that there are exactly 2^i input links from which ℓ can be reached and exactly 2^{k-i} output links that can be reached from ℓ. Because the configuration is α-bounded,

$$\lambda_\ell(C) \leq \alpha 2^{-i} \quad \text{and} \quad \lambda_\ell \leq \alpha 2^{k-i}$$

So,

$$\lambda_\ell(C) \leq \alpha \min_{0 \leq i \leq k} \{2^i, 2^{k-i}\} \leq \alpha 2^{\lfloor k/2 \rfloor} \leq \alpha\sqrt{n} \qquad \Box$$

The bound in Theorem 2.1 can be achieved as illustrated in Figure 3. We note that delta networks are readily generalized to networks in which each internal node has m input ports and m output ports. The following theorem is proved in the same way as Theorem 2.1

THEOREM 2.2. *Let $C = \{c_1, \ldots, c_r\}$ be an α-bounded configuration for an n-port m-ary delta network. Then $\lambda_\ell(C) \leq \alpha\sqrt{n}$ for all links ℓ.*

The bounds in the above theorems and the fact that there are traffic patterns that achieve the bounds, lead to the conclusion that the binary and m-ary delta networks can perform poorly under worst-case loading conditions. This has been observed previously and various approaches have been proposed to remedy the situation. We analyze two such approaches here. The first is to add one or more stages of distribution nodes at the front of a delta network.

We denote a delta network with k routing stages and d distribution stages as $D_{k,d}$, which we define by $D_{k,d} = D_d \Box D_{k-d} \Box D_d$. This is illustrated in Figure 4.
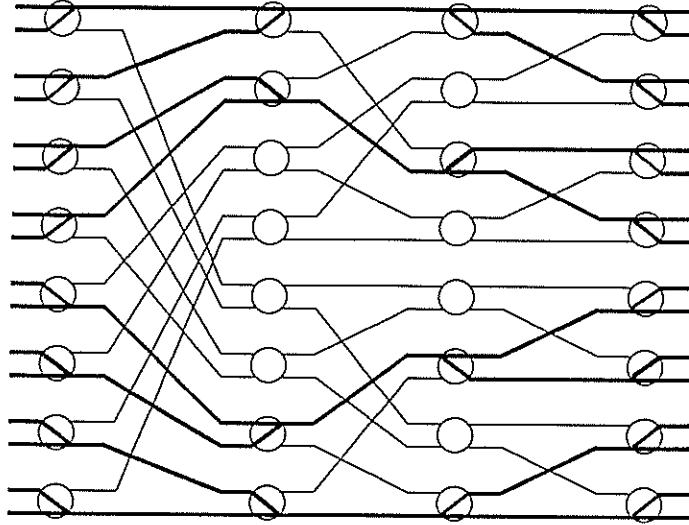
Connections $(i, i, 1)$ for $0 \leq i \leq 15$.

Figure 3: Worst-Case Configuration for a Delta Network

If we consider the load induced by a connection (x, y, ρ) on the links in such a network, we note that for any node u in the first d stages that lies on a path from x to y, both of u's output links lie on paths from x to y, hence the incoming load from the connection is distributed across u's output links. In contrast, any node v in the last k stages is on at most one path from x to y. We refer to the nodes in the first d stages as distribution nodes and the nodes in the last k stages as routing nodes.

THEOREM 2.3. *Let* $C = \{c_1, \ldots, c_r\}$ *be an* α*-bounded configuration for* $D_{k,d}$. *Then* $\lambda_\ell(C) \leq \alpha n 2^{-\lceil (k+d)/2 \rceil}$ *for all links* ℓ.

Proof. For any link ℓ, we define $C_\ell \subseteq C$ to be the set of connections $(x, y, \rho) \in C$ for which ℓ lies on some path from x to y. If ℓ is in stage i and $c_h = (x_h, y_h, \rho_h)$ is any connection, then

$$\lambda_\ell(c_h) \leq \begin{cases} \rho_h 2^{-i} & 0 \leq i \leq d \\ \rho_h 2^{-d} & d \leq i \leq k \\ \rho_h 2^{-(k+d-i)} & k \leq i \leq k+d \end{cases}$$

Also, because the configuration is assumed to be α-bounded

$$\sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^i \quad \text{and} \quad \sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^{k+d-i}$$
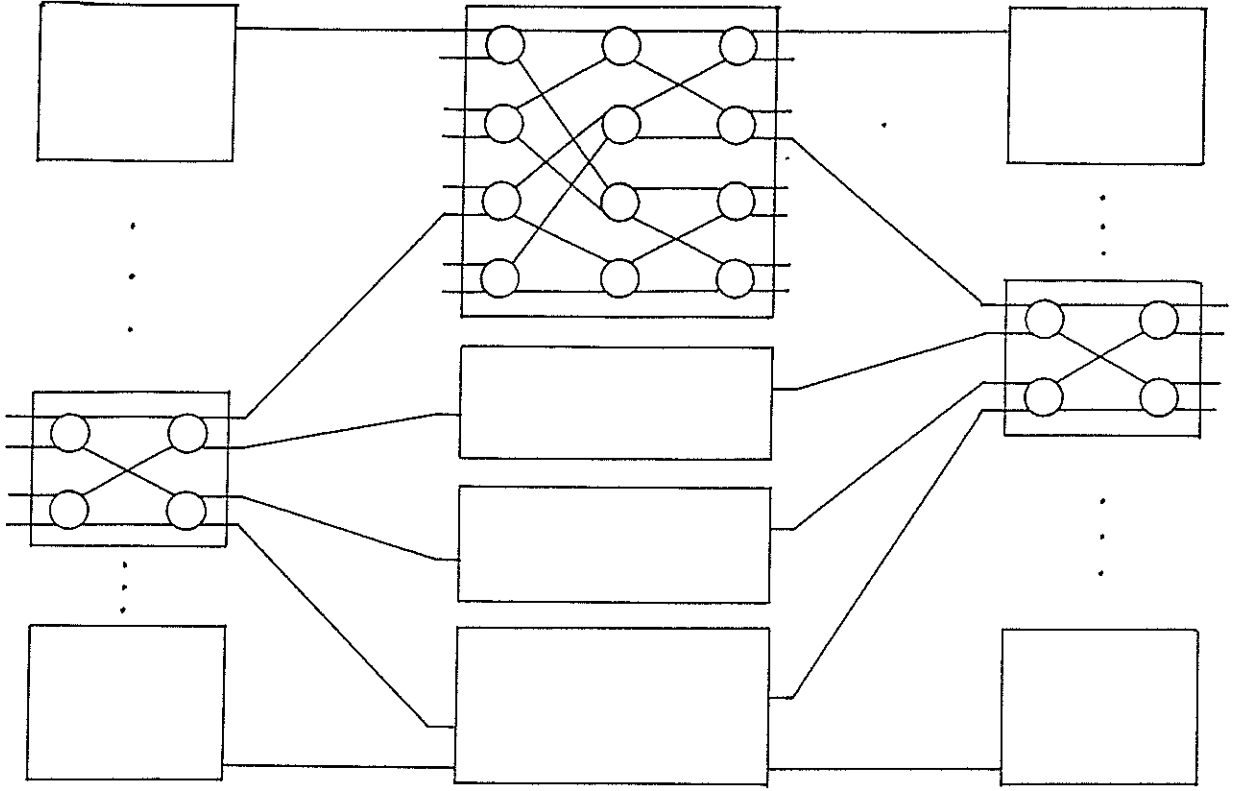
Figure 4: Construction of Delta Network with Distribution Stages

So, if $0 \leq i \leq d$,

$$\lambda_\ell(C) = \sum_{c_h \in C_\ell} \lambda_\ell(c_h) \leq 2^{-i} \sum_{c_h \in C_\ell} \rho_h \leq \alpha$$

If $k \leq i \leq k + d$,

$$\lambda_\ell(C) = \sum_{c_h \in C_\ell} \lambda_\ell(c_h) \leq 2^{-(k+d-i)} \sum_{c_h \in C_\ell} \rho_h \leq \alpha$$

And if $d \leq i \leq k$,

$$\begin{aligned}
\lambda_\ell(C) &\leq 2^{-d} \sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^{-d} \min_{d \leq i \leq k} \left\{ 2^i, 2^{k+d-i} \right\} \\
&= \alpha 2^k \min_{d \leq i \leq k} \left\{ 2^{-(k+d-i)}, 2^{-i} \right\} \leq \alpha n 2^{-\lceil (k+d)/2 \rceil} \quad \square
\end{aligned}$$

The bound in Theorem 2.3 is the best possible, as illustrated by the example in Figure 5. Theorem 2.3 tells us that every time we add two distribution stages, we reduce the worst-case load by a factor of 2. To achieve a robust network, we require $d = k - 1$. Also note that with respect to worst-case loading, it never makes sense to have $k + d$ an even number, since a network with one fewer distribution stage has the same worst-case loading characteristics.
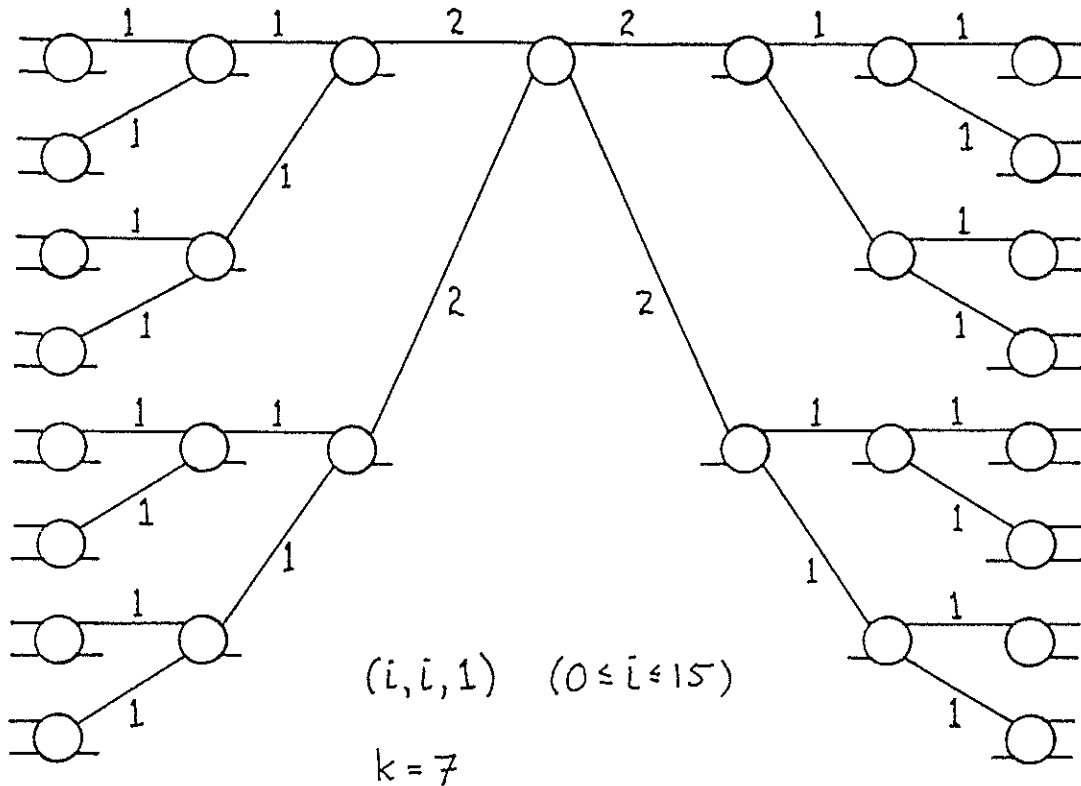
Figure 5: Worst-Case Configuration for a Delta Network with Distribution Stages

We note that this result can be readily generalized to networks with nodes having m input and output ports. The bound in the statement of the theorem becomes $nm^{-\lceil(k+d)/2\rceil}$ (with $k = \log_m n$) and the proof is essentially the same as the one given.

In [16], Lea proposes a variant of the delta network that we refer to as the *alternate routing network*. We can define this network recursively using the composition operation. The base network is denoted by A_1 and consists of four input ports and four output ports connected to a single internal node. For $i > 1$, $A_i = A_1 \boxempty A_{i-1}$. An example of an alternate routing network is given in Figure 6. Note that an alternate routing network with k stages has $n = 2^{k+1}$ ports. Given any connection (x, y, ρ), if u is in the first $k - 1$ stages and lies on some path from x to y, then two of u's output links lie on paths from x to y. Consequently, whatever load is induced on the input links of u will be shared by two of u's output links. We now characterize the worst-case load for such a network. The following theorem characterizes the worst-case loading of an alternate routing fabric. We note that essentially the same result is stated (in somewhat different terms) in [16].

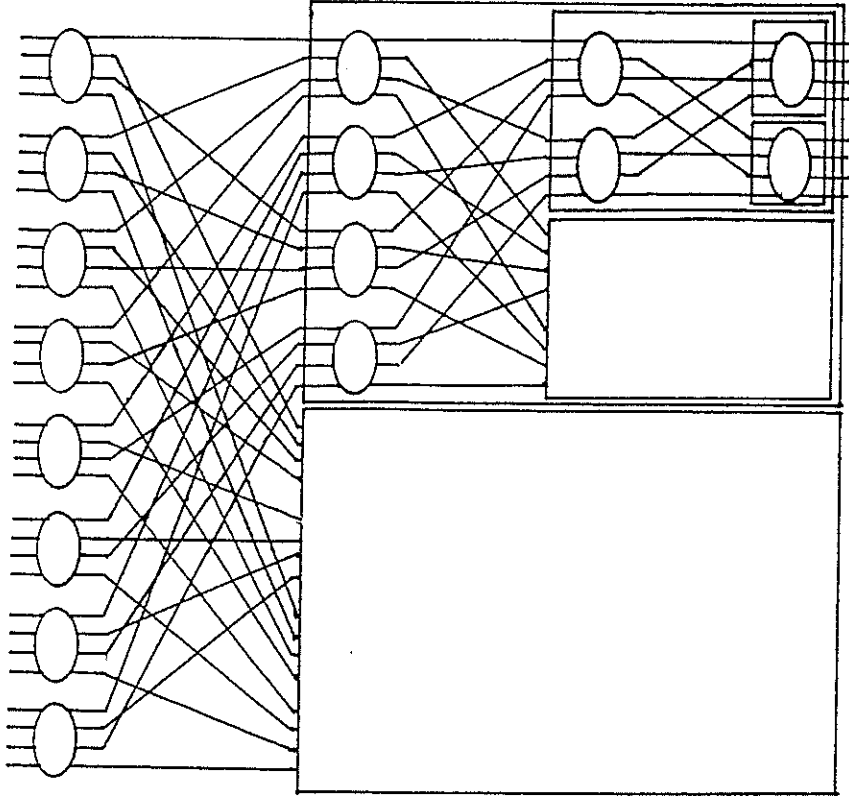THEOREM 2.4. *Let* $C = \{c_1, \ldots, c_r\}$ *be an* α-*bounded configuration for* A_k.

Figure 6: Construction of Alternate Routing Network

Then $\lambda_\ell(C) \leq \alpha n^{1/3}$ for all links ℓ.

Proof. For any link ℓ, we define $C_\ell \subseteq C$ as in the proof of Theorem 2.3. If ℓ is in stage i and $c_h = (x_h, y_h, \rho_h)$ is any connection, then

$$\lambda_\ell(c_h) \leq \begin{cases} \rho_h 2^{-i} & 0 \leq i \leq \lfloor k/2 \rfloor \\ \rho_h 2^{-(k-i+1)} & \lfloor k/2 \rfloor + 1 \leq i \leq k-1 \end{cases}$$

Also, because the configuration is assumed to be α-bounded

$$\sum_{c_h \in C_\ell} \rho_h \leq \alpha 4^i \quad \text{and} \quad \sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^{k-i+1}$$

So, if $i \leq \lfloor k/3 \rfloor$,

$$\lambda_\ell(C) = \sum_{c_h \in C_\ell} \lambda_\ell(c_h) \leq 2^{-i} \sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^i \leq \alpha 2^{\lfloor k/3 \rfloor} \leq \alpha n^{1/3}$$

If $\lfloor k/3 \rfloor + 1 \leq i \leq \lfloor k/2 \rfloor$

$$\lambda_\ell(C) \leq 2^{-i} \sum_{c_h \in C_\ell} \rho_h \leq \alpha 2^{k-2i+1} \leq \alpha n^{1/3}$$
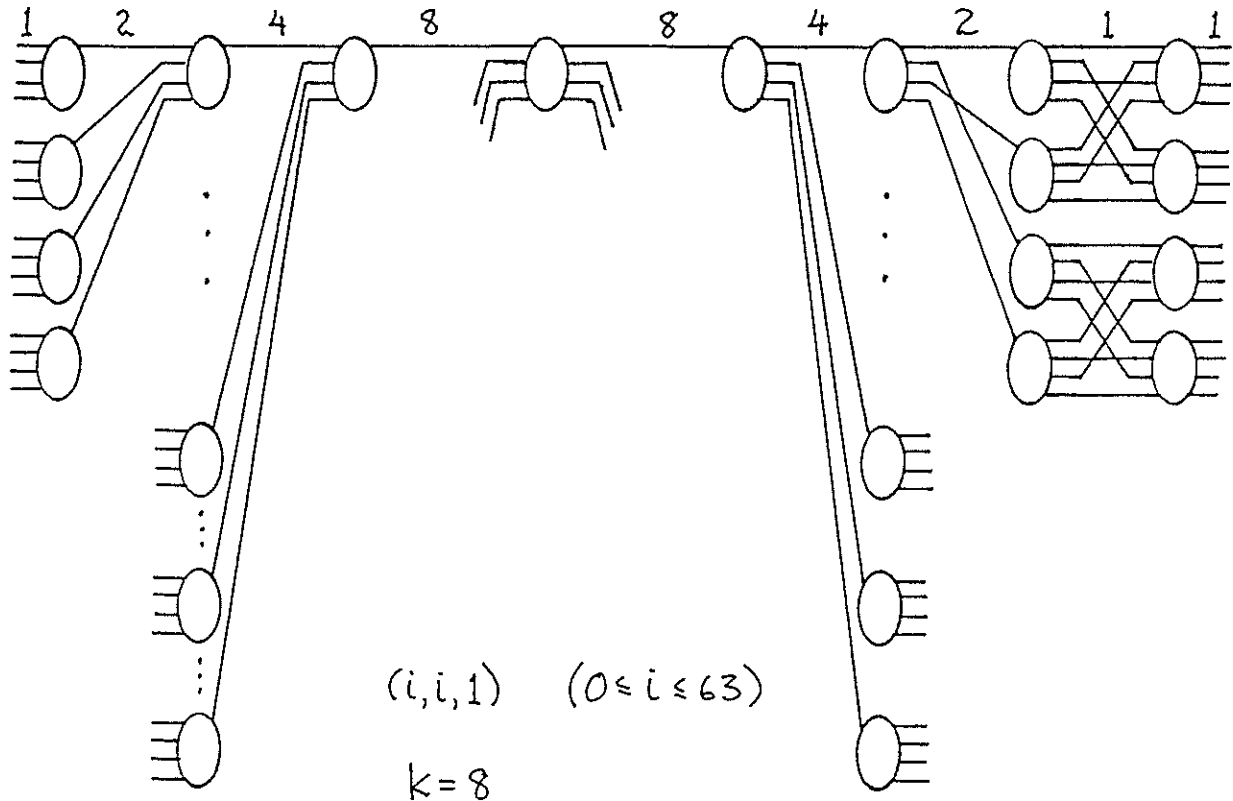
Figure 7: Worst-Case Configuration for Alternate Routing Network

And if $\lfloor k/2 \rfloor + 1 \le i$

$$\lambda_\ell(C) \le 2^{-(k-i+1)} \sum_{c_h \in C_\ell} \rho_h \le \alpha \qquad\qquad \square$$

3. Loading in Copy Networks

Recently, several research groups have been studying the problem of multipoint communication in high speed packet networks and have developed new types of switching systems to support such networks. In [21], this author describes a particular switching system comprising two constituent networks; a *copy network* used to replicate packets belonging to multipoint connections and a *routing network* used to route the resulting copies to the appropriate outgoing links. In this section we study the worst-case loading in the copy network of [21], as well as several variants.

When dealing with copy networks, we must modify our definition of connection. In the current context, we define a connection to be an an ordered triple (x, F, ρ), where x is the input port of the copy network where packets belonging to the connection enter, F is the *fanout* of the connection and ρ is the load
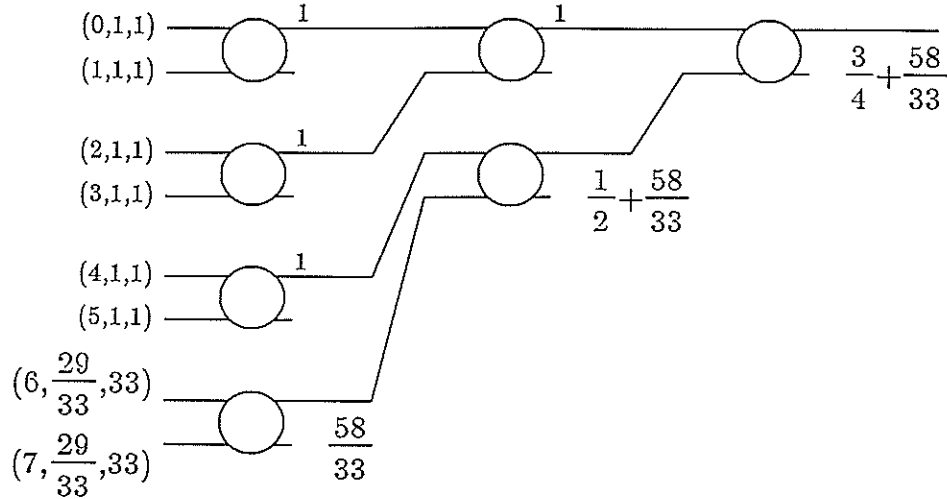
Figure 8: Worst-Case Configuration for Copy Network

induced by the connection at the input port x. The fanout of the connection is the number of copies that must be produced by the copy network for each input packet. We say that a traffic configuration C is α-bounded if $\lambda_\ell(C) \le \alpha$ for all input ports ℓ and $\sum_{(x,F,\rho)\in C} \rho F \le \alpha n$, where n is the number of input and output ports. A legal configuration is one that is 1-bounded.

Reference [21], describes a copy network that is topologically identical to a delta network. However, the nodes of a copy network replicate received packets under certain conditions. Specifically, a node may replicate a packet if the number of output ports reachable from that node is less than $2F$, where F is the fanout of the connection the packet belongs to. Packets that are not replicated are routed to an arbitrarily selected output. Hence, if α is the load induced on the input links of a node u by a connection (x, F, ρ), then the load induced on each of u's output links is α, if the number of output ports that can be reached from u is $< 2F$ and $\alpha/2$ otherwise.

Given these definitions, we find that for a connection $c = (x, F, \rho)$ and a link ℓ in stage i,

$$\lambda_\ell(c) = \begin{cases} 0 & \text{if there is no path from input } x \text{ to link } \ell \\ \rho 2^{-i} & \text{if there is a path and } 0 \le i \le k - \lceil \log_2 F \rceil \\ \rho 2^{-(k - \lceil \log_2 F \rceil)} & \text{if there is a path and } i \ge k - \lceil \log_2 F \rceil \end{cases}$$

Our first theorem, which was first proved in [2] shows that the worst-case loading in a copy network is bounded.

THEOREM 3.1. *Let* $C = (c_1, \ldots, c_r)$ *be any α-bounded configuration for an n-port copy network. Then,* $\lambda_\ell(C) \le 3\alpha$ *for all links ℓ.*

Proof. Let $k = \log_2 n$ be the number of stages in the network. For any link ℓ, we define $C_\ell \subseteq C$ to be the set of connections $(x, F, \rho) \in C$ for which ℓ lies on some path from x to y. If ℓ is in stage i and $c_h = (x_h, F_h, \rho_h)$ is any connection and $f_h = \lceil \log_2 F_h \rceil$, then

$$\lambda_\ell(c_h) \leq \begin{cases} \rho_h 2^{-i} & 0 \leq i < k - f_h \\ \rho_h 2^{-(k-f_h)} & k - f_h \leq i \leq k \end{cases}$$

Let

$$\begin{aligned} C_\ell^< &= \{c_h \in C_\ell \,|\, i < k - f_h\} \\ C_\ell^\geq &= \{c_h \in C_\ell \,|\, i \geq k - f_h\} \end{aligned}$$

Because the configuration is assumed to be α-bounded, $\sum_{c_h \in C_\ell^<} \rho_h \leq \alpha 2^i$ So,

$$\lambda_\ell(C_\ell^<) = \sum_{c_h \in C_\ell^<} \lambda_\ell(c_h) \leq 2^{-i} \sum_{c_h \in C_\ell^<} \rho_h \leq \alpha$$

Also,

$$\lambda_\ell(C_\ell^\geq) = \sum_{c_h \in C_\ell^\geq} \lambda_\ell(c_h) \leq 2^{-k} \sum_{c_h \in C_\ell^\geq} 2^{f_h} \rho_h < (2/n) \sum_{c_h \in C_\ell^\geq} F_h \rho_h \leq 2\alpha$$

Since $\lambda_\ell(C) = \lambda_\ell(C_\ell^<) + \lambda_\ell(C_\ell^\geq)$, $\lambda_\ell(C) < 3\alpha$. $\square$

The bound in Theorem 3.1 cannot be improved. Figure 8 shows an example of a bad traffic configuration for a six stage copy network. This example can be generalized to yield traffic configurations that induce loads approaching 3 on some internal links.

Copy networks can also be constructed using nodes with $m > 2$ input and output ports. In such networks, a node replicates a packet m times if the number of reachable output ports is less than mF. Surprisingly, the worst-case performance of such a copy network is worse than for a copy network constructed from binary nodes.

THEOREM 3.2. *Let $C = (c_1, \ldots, c_h)$ be an α-bounded configuration for an n-port copy network constructed from m-port nodes. Then $\lambda_\ell(C) < \alpha(m + 1)$, for all links ℓ.*

The proof of this is very similar to that of Theorem 3.1. The bound is the best possible; it is not difficult to construct legal traffic configurations that induce loads approaching $m + 1$ on some internal links.

As with routing networks, we can improve the worst-case performance of a copy network by adding distribution stages. The topology of such a network is identical to a routing network with added distribution stages. The effect on the worst-case loading is captured by the following theorem.

THEOREM 3.3. *Let $C = (c_1, \ldots, c_h)$ be an α-bounded configuration for a copy network with k copy stages and d distribution stages. Then $\lambda_\ell(C) < \alpha(1 + 2^{1-d})$, for all links ℓ in stages 0 to $k + d - 1$; $\lambda_\ell < 2\alpha$ for all links ℓ in stage $k + d$.*

Proof. For any link ℓ, we define $C_\ell \subseteq C$ as in the proof of Theorem 3.1. Let $c_h = (x_h, F_h, \rho_h)$ denote an arbitrary connection in C, $f_h = \lceil \log_2 F_h \rceil$ and let ℓ be any link in stage i. We define

$$C_\ell^< = \{c_h \in C_\ell \mid i < k + d - f_h\}$$
$$C_\ell^\geq = \{c_h \in C_\ell \mid i \geq k + d - f_h\}$$

If $i \leq k$, then

$$\lambda_\ell(c_h) \leq \begin{cases} \rho_h 2^{-i} & i < k + d - f_h \\ \rho_h 2^{-(k+d-f_h)} & i \geq k + d - f_h \end{cases}$$

and $\sum_{c_h \in C_\ell^<} \rho_h \leq \alpha 2^i$. So,

$$\lambda_\ell(C) = \lambda_\ell(C_\ell^<) + \lambda_\ell(C_\ell^\geq) \leq 2^{-i} \sum_{c_h \in C_\ell^<} \rho_h + 2^{-(k+d)} \sum_{c_h \in C_\ell^\geq} 2^{f_h} \rho_h$$
$$\leq \alpha + (2^{1-d}/n) \sum_{c_h \in C_\ell^\geq} F_h \rho_h \leq \alpha(1 + 2^{1-d})$$

For $k < i \leq k + d$,

$$\lambda_\ell(c_h) \leq \begin{cases} \rho_h/n & i < k + d - f_h \\ (\rho_h/n) 2^{i-(k+d-f_h)} & i \geq k + d - f_h \end{cases}$$

So,

$$\lambda_\ell(C) = \lambda_\ell(C_\ell^<) + \lambda_\ell(C_\ell^\geq) \leq (1/n) \sum_{c_h \in C_\ell^<} \rho_h + \left(2^{-(k+d-i)}/n\right) \sum_{c_h \in C_\ell^\geq} 2^{f_h} \rho_h$$
$$< (1/n) \left[\sum_{c_h \in C_\ell^<} \rho_h + \sum_{c_h \in C_\ell^\geq} F_h \rho_h \right] + \left((2^{1-(k+d-i)} - 1)/n\right) \sum_{c_h \in C_\ell^\geq} F_h \rho_h$$
$$\leq \alpha + \left((2^{1-(k+d-i)} - 1)/n\right) \sum_{c_h \in C_\ell^\geq} F_h \rho_h \leq \alpha \left(1 + (2^{1-(k+d-i)} - 1)\right)$$

This last expression is $\leq 2\alpha$ if $i = k + d$ and is $\leq \alpha$ if $i < k + d$. $\square$

Theorem 3.3 shows that the worst-case loading in a copy network can be brought very close to α in all but the last stage links, by adding a few distribution stages. Finally, we note that Theorem 3.3 can be generalized to copy networks constructed with m-ary nodes. In this case the bound on the worst-case loading becomes $\alpha(1 + m^{1-d})$ for all but the last stage and m for the last stage.

4. Closing Remarks

The fluid flow loading analysis method described here provides a useful first-order characterization of a large class of packet switching fabrics. The method is fairly easy to apply and leads to useful insights that can guide the switching system architect to better designs. It is not a complete characterization, as it ignores queueing and contention, but when used in conjunction with queueing and simulation models based on uniform random traffic, it can provide the designer and performance analyst with a more complete understanding of system performance.

The specific results obtained here are of substantial interest. In particular, the analysis of the effect of distribution stages for both point-to-point and copy networks quantifies the improvement obtained by adding distribution stages and allows the system architect to make intelligent choices when making the trade-off between performance and cost. We have found that the use of nodes with more than two ports does not improve the worst-case loading of point-to-point delta networks, either with or without distribution and surprisingly that it leads to worse loading in copy networks.

There are several open problems that remain to be addressed. One fundamental question is whether or not there is a lower bound on the number of m input, m output nodes in a point-to-point switching network that is robust. This is similar to the question of the minimum number of crosspoints in a non-blocking network, but cannot be deduced directly, since packet switching fabrics are allowed the freedom of splitting the traffic from individual connections across multiple paths. A related question is whether or not there exist robust fabrics with fewer nodes than the ones described here. Our results for copy networks have not led to a robust network because of the excess loading in the last stage; the only apparent remedy for this at the moment is to double the number of nodes in the last stage, with each node having just one input link but two output links. This of course doubles the number of copy network outputs which makes it a rather costly solution, but does give a robust network. Another problem of interest is to apply our method of loading analysis to other networks that have been proposed in the literature.

Finally, we note that there remains a need for a more comprehensive approach to performance modelling for these types of switching networks; one that takes into account both the effect of user connections and queueing. In addition, we need more realistic models of the packet arrival processes associated with each connection and the resulting performance impact. The common assumption of Poisson arrivals is grossly inaccurate for many of the applications that fast packet switching is intended to support. These applications typically alternate between active states in which they transmit data at high rates and inactive states in which

they transmit little or no data. Current performance models do not account for this behavior and may lead to incorrect conclusions.

References

[1] Batcher, K. E. "Sorting Networks and Their Applications," *Spring Joint Computer Conference*, 1968, 307–314.

[2] Bubenik, Richard. "Performance Evaluation of a Broadcast Packet Switch," Washington University Computer Science Department, MS thesis, 8/85.

[3] Bubenik, Richard and Jonathan S. Turner. "Performance of a Broadcast Packet Switch." *International Communications Conference*, 6/87.

[4] Coudreuse, J. P. and M. Servel. "Asynchronous Time-Division Techniques: An Experimental Packet Network Integrating Videocommunication," *International Switching Symposium*, 1984.

[5] Coudreuse, J. P. and M. Servel. "Prelude: An Asynchronous Time-Division Switched Networ,," *International Communications Conference*, 1987.

[6] De Prycker, M., and J. Bauwens. "A Switching Exchange for an Asynchronous Time Division Based Network," *International Communications Conference*, 1987.

[7] Dias, D. M. and J. R. Jump, "Packet Switching Interconnection Networks For Modular Systems," *Computer*, vol. 14, no. 12, 12/81, 43-53

[8] Dias, D. M. and J. R. Jump, "Analysis and Simulation of Buffered Delta Networks," *IEEE Transactions on Computers*, vol. C-30, no. 4, 4/81, 273-282

[9] Dias, D. M. and Manoj Kumar. "Packet Switching in $N \log N$ Multistage Networks," *Globecom*, 12/84, 114–120.

[10] Feng, Tse-yun. "A Survey of Interconnection Networks," *Computer*, vol. 14, no. 12, 12/83, 12–30.

[11] Giorcelli, C. Demichelis, G. Giandonato and R. Melen. "Experimenting with Fast Packet Switching Techniques in a First Generation ISDN Environment," *International Switching Symposium*, 3/87.

[12] Goke, L. R. and G. J. Lipovski, "Banyan Networks for Partitioning Multiprocessor Systems," *Symposium on Computer Architecture*, 4/79, 182-187.

[13] Hayward, G, L. Linnell, D. Mahoney and L. Smoot. "A Broadband ISDN Local Access System Using Emerging Technology Components," *International Switching Symposium*, 3/87.

[14] Jenq, Yih-Chyun. "Performance Analysis of a Packet Switch Based on a Single-Buffered Banyan Network," *IEEE Journal on Selected Areas in Communications*, vol. SAC-1, no. 6, 12/83, 1014–1021.

[15] Kruskal, C. P. and M. Snir. "The Performance of Multistage Interconnection Networks for Multiprocessors," *IEEE Transactions on Computers*, vol. C-32, no. 12, 12/83, 1091–1098.

[16] Lea, Chin-tau. "The Load-Sharing Banyan Network," *IEEE Transactions on Computers*, 12/86.

[17] Kulzer, John J. and Warren A. Montgomery. "Statistical Switching Architectures for Future Services," *International Switching Symposium*, 5/84.

[18] Patel, J. H. "Performance of Processor-Memory Interconnections for Multiprocessors," *IEEE Transactions on Computers*, vol. C-30, no. 10, 10/81, 771-780.

[19] Takeuchi, T., H. Suzuki, S. Hayano, H. Niwa and T. Yamaguchi, "An Experimental Synchronous Composite Packet Switching system," *Zurich Seminar on Digital Communication*, 3/86.

[20] Turner, Jonathan S. and L. F. Wyatt, "Alternate Paths in a Self-Routing Packet Switching Network," United States Patent #4,550,397, 10/29/85.

[21] Turner, Jonathan S. "Design of a Broadcast Packet Network," *Infocom*, 4/86.